

# Messung experimenteller Problemlöse- und Diagrammkompetenzen und deren Zusammenhang mit kognitiven Fähigkeiten im Fach Naturwissenschaften im 5. Jahrgang einer Integrierten Gesamtschule

Kerstin Höner, Kristiena Matis & Deborah Ammon

Technische Universität Braunschweig, Institut für Fachdidaktik der Naturwissenschaften, Abteilung Chemie und Chemiedidaktik

## Kurzfassung

Experimentelle Problemlösekompetenzen sowie auch Kompetenzen beim Umgang mit Diagrammen sind wichtige Ziele des naturwissenschaftlichen Unterrichts. Eng verbunden mit der Zielsetzung des Aufbaus entsprechender Schülerfähigkeiten ist die Frage der Diagnostik.

In der vorliegenden Studie wird als Grundlage für die Diagnose der oben genannten Fähigkeiten eine naturwissenschaftlich-experimentelle Aufgabensequenz verwendet, die auch die eigenständige Konstruktion eines Diagramms beinhaltet.

Da viele Studien gezeigt haben, dass die Ergebnisse zwischen den Leistungen schriftlicher Tests und denen beim realen Experimentieren nur gering miteinander korrelieren, wurden die SchülerInnen hier bei der Durchführung des Realexperiments videografiert und die Daten qualitativ analysiert. Parallel dazu erstellten die SchülerInnen auch schriftliche Protokolle, die quantitativ ausgewertet wurden. Da aus theoretischer Perspektive angenommen werden kann, dass kognitive Fähigkeiten einen Einfluss haben, wurden diese als Kovariablen mit erhoben.

Es zeigten sich positive, schwach signifikante Korrelationen zwischen den Leistungen bei der schriftlichen Dokumentation zum Experimentieren und den kognitiven Fähigkeiten der SchülerInnen. Die Kompetenz hinsichtlich der Diagrammkonstruktion ist insgesamt eher schwach ausgeprägt und korreliert nicht signifikant mit den kognitiven Fähigkeiten.

Des Weiteren konnten mit der qualitativen Inhaltsanalyse aus den Videografiedaten Faktoren identifiziert werden, die den Erfolg der SchülerInnen beim experimentellen Problemlösen beeinflussen können. Ein Vergleich der quantitativen und qualitativen Daten zeigt, dass die schriftlichen Aufzeichnungen der SchülerInnen keine differenzierte individuelle Diagnose hinsichtlich ihrer experimentellen Problemlösefähigkeiten erlaubt, wenn die SchülerInnen in Gruppen arbeiten.

*Schlüsselwörter:* naturwissenschaftlich-experimentelles Problemlösen; kognitive Fähigkeiten; Diagrammkompetenz; Videografie

## Measuring experimental problem-solving and graph competencies and their relationships to cognitive abilities in a 5th year science class at a - comprehensive school (Integrierte Gesamtschule)

### Summary

Experimental problem-solving competencies as well as competencies in dealing with graphs are important goals of science education. Closely linked to the objective of building up the respective skills is the question of diagnostics.

The study presented here uses an experimental science task sequence, which also includes the independent construction of a graph, as a basis for diagnosing the above-named skills.

Since many studies have confirmed that pupils' performance in paper-pencil tests correlates only slightly with their performance in practical experiments, the pupils were videotaped while conducting the experiment and the data were analysed qualitatively. Simultaneously the pupils also

created lab records that were analysed quantitatively. Since it can be assumed from a theoretical perspective that cognitive abilities have an influence, they were included as covariates.

Positive, weakly significant correlations were found between the performance in the written documentation of the experiment and cognitive abilities of the pupils. The competence of plotting a graph is generally rather weak and does not correlate significantly with cognitive abilities.

Furthermore, qualitative content analysis of the videotaped data reveals factors that may influence the pupils' performance in experimental problem-solving. A comparison of the quantitative and qualitative data shows that the pupils' written documentations do not allow a differentiated individual diagnosis with regard to their experimental problem-solving skills when the pupils work in groups.

*Key words: experimental problem-solving; cognitive abilities; graph competence; videography*

---

## **Einleitung**

Problemlösekompetenzen gehören zu den grundlegenden Schlüsselqualifikationen, die für die Bewältigung sowohl schulischer als auch lebensweltlicher und beruflicher Anforderungen bedeutsam sind. Die Problemlösekompetenz deutscher SchülerInnen liegen den Ergebnissen der PISA-Studie von 2012 zufolge zwar über dem OECD-Durchschnitt, sollten aber aufgrund der gestiegenen beruflichen und gesellschaftlichen Anforderungen weiter gefördert werden (Zentrum für Internationale Vergleichsstudien 2013). Dies trifft auch auf mathematische Darstellungen zu, die im naturwissenschaftlichen Unterricht ebenfalls eine wichtige Rolle spielen (Lachmayer 2008).

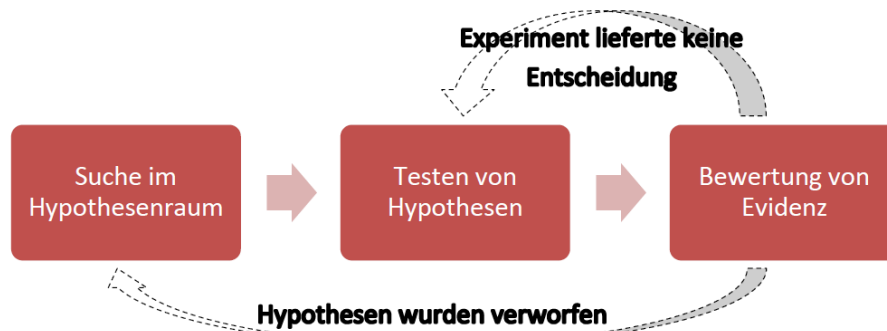
Die seit den Arbeiten von Dörner (1976) als eigenständige Fähigkeit angesehene komplexe Problemlösekompetenz basiert darauf, dass kaum fachspezifische, sondern vorwiegend allgemeine, interdisziplinäre und vorwissensneutrale Inhalte, also prozessbezogene Kompetenzen, entscheidend sind. Trotzdem sind Konzepte der Wissensaneignung und Wissensanwendung, die für das komplexe Problemlösen wichtig sind, auch für das schulische Lernen im Hinblick auf Kompetenzentwicklung bedeutsam (Kretzschmar, Neubert, Greiff 2014). So setzt sich das schulische Problemlösen oftmals aus einem kognitiven, bereichsübergreifenden Anteil, der zum Beispiel die Wissensgenerierung durch Wechselwirkung mit dem System oder die Reflexion des Problemlöseprozesses beinhaltet und einem spezifischen Anteil, der curriculare Kompetenzen der einzelnen Schulfächer umfasst, zusammen (Scherer 2014a).

## **Naturwissenschaftlich-experimentelle Problemlösekompetenz und deren Modellierung**

Der Begriff „Problemlösekompetenz“ bezeichnet die individuelle Fähigkeit, Probleme aus verschiedenen Bereichen zu verstehen und zu lösen. Als „Problemlösen“ wird der Prozess bezeichnet, der sich mit der Beseitigung eines Problems befasst, für dessen Bewältigung zunächst keine Strategie bzw. keine Routinelösung vorhanden ist (Dörner 1976; Betsch, Funke, Plessner 2011). Das Problem stellt eine Barriere dar, die zwischen einem Ausgangszustand und einem zu erreichenden Zielzustand steht (Becker-Carus, Wendt 2017.) Sobald die Problemlösenden eine routinierte Vorgehensweise zum Erreichen des Zielzustands entwickelt haben, wird das frühere „Problem“ zu einer herkömmlichen „Aufgabe“. Folglich ist es von den zum jeweiligen Zeitpunkt vorhandenen individuellen Fähigkeiten und vom Vorwissen der Problemlösenden abhängig, ob eine Herausforderung als „Aufgabe“ oder „Problem“ klassifiziert wird (Dörner 1976).

Die naturwissenschaftlich-experimentelle Problemlösekompetenz bezieht sich auf Probleme, die im Sinne der Erkenntnisgewinnung in den Naturwissenschaften zu lösen sind und besteht aus mehreren Phasen. Um diese Kompetenz diagnostizierbar zu machen, wurden Kompetenzmodelle entwickelt, die verschiedene Teilkompetenzen umfassen. Je nach Ansatz gibt es jedoch konzeptuelle Trennungen oder Zusammenlegungen einzelner Elemente, sodass die Befunde zur Dimensionierung dieser Teilkomponenten bisher inkonsistent sind (Vorholzer, Hägele, von Aufschnaiter 2017, 27).

Ein häufig verwendetes Modell zur Beschreibung der Erkenntnisprozesse beim naturwissenschaftlichen Problemlösen ist das *Scientific Discovery as Dual Search-Modell* (SDDS-Modell) (Klahr, Dunbar 1988; Hammann 2007). Das SDDS-Modell beschreibt das Experimentieren als eine Suche in einem Hypothesen- und einem Experimentierraum sowie die Bewertung von Evidenz (s. Abbildung 1).



**Abbildung 1:** SDDS-Modell nach Klahr (2000). (Wenzel 2016, S. 9)

Bei der Suche im Hypothesenraum wird eine überprüfbare, plausible Hypothese entwickelt. Es folgt das Testen der Hypothese im Experimentierraum. Durch die Auswahl einer geeigneten, experimentellen Vorgehensweise werden die Ergebnisse erhalten, mit denen dann die Hypothese überprüft wird. In der dritten Hauptkomponente, der Bewertung von Evidenzen, wird auf Grundlage der Erkenntnisse aus der Phase des Testens von Hypothesen entschieden, ob die eingangs gebildete Hypothese beibehalten oder verworfen wird (Hammann 2007). Sollte die Hypothese verworfen werden, ist eine erneute Suche im Hypothesenraum nötig. Wenn die bisher erhaltenen experimentellen Erkenntnisse nicht ausreichen, um zu entscheiden, ob die Hypothese beibehalten oder verworfen werden soll, sind weitere Untersuchungen in der Komponente des Testens von Hypothesen notwendig, was auch eine erneute Suche im Experimentierraum einschließt (Wenzel 2016).

### **Messung naturwissenschaftlich-experimenteller Problemlösekompetenz und kognitiver Fähigkeiten**

Das naturwissenschaftlich-experimentelle Problemlösen im Sinne der Erkenntnisgewinnung beinhaltet in allen Modellen das Bilden von Hypothesen, die Auswahl geeigneter Experimente zur Überprüfung der Hypothesen sowie das Überprüfen der Hypothesen anhand der erhaltenen experimentellen Ergebnisse. Diese Kompetenzen sollten von den jeweiligen Testverfahren erfasst werden, indem sie durch verschiedene Indikatoren, jeweils abhängig von der Klassenstufe und vom Experiment, operationalisiert werden (Schreiber, Theyßen, Schecker 2009). Bei der Kompetenzerfassung anhand von Realexperimenten sind neben der quantitativen Auswertung von Protokollen auch eine gezielte, direkte Beobachtung und Videografie mit anschließender kriteriengeleiteter Auswertung möglich (Baur 2015).

Wenn auch für die Alltagspraxis in der Regel zu aufwändig, bietet die Videografie von SchülerInnen beim realen Experimentieren für Forschungszwecke viele Vorteile, weil die Aufzeichnungen wiederholt angeschaut und analysiert werden können, sodass auf verschiedene Fragestellungen und Perspektiven fokussiert werden kann. Außerdem können die Videoaufnahmen von verschiedenen Personen betrachtet werden, um so die Auswertungsobjektivität sicherzustellen. Darüber hinaus ist es möglich, die sprachliche Kommunikation der SchülerInnen untereinander zu transkribieren, um sie tiefergehend inhaltsanalytisch auszuwerten (Lotz 2016).

Bei vielen Problemen kann fachspezifisches Vorwissen einen großen Einfluss auf die Problemlösung haben, sodass es für die Diagnose allgemeiner naturwissenschaftlich-experimenteller Problemlösekompetenz sinnvoll ist, weitgehend vorwissensneutrale Probleme zu verwenden, damit

das Vorwissen kein schwierigkeiterzeugendes Merkmal ist (Vorholzer, Hägele, von Aufschnaiter 2017; Messick 1995). Eine weitere Möglichkeit ist es, im Vorfeld der Untersuchung das Vorwissen der SchülerInnen möglichst anzugleichen, um zu vermeiden, dass einzelne Personen mit fachspezifischem Vorwissen bevorteilt werden (Greiff, Fischer 2013).

Ein entscheidendes Merkmal des Problemlösens ist das Überwinden von Barrieren, was immer mit kognitiven Anstrengungen verbunden ist (Dörner 1976). Aus theoretischer Perspektive ist deshalb anzunehmen, dass es einen Zusammenhang zwischen der Problemlösekompetenz und den kognitiven Fähigkeiten des Problemlösers gibt (s. auch Nehring 2014, 96). Die empirische Befundlage zu dieser Relation ist aber divers (Greiff 2012; Leutner et al. 2005; Leutner et al. 2012; Süß 1996; Wirth, Leutner, Klieme 2005) und hängt von der Art der Problemstellung ab.

Ein häufig verwendeter standardisierter Test für die Messung allgemeiner kognitiver Fähigkeiten für die Jahrgangsstufen 4 bis 12 ist der *KFT 4-12+R* von Heller und Perleth (2000). Dieser enthält drei Testteile mit jeweils drei Subtests. Der Test ist als „Power-Speed-Test“ angelegt, bei dem die Zeit für die Bearbeitungszeit begrenzt ist (Kipman, Kohlböck, Weigluny 2012). Der Test ist vielfach erprobt, so dass die Einhaltung der Gütekriterien bei vorschriftsmäßiger Anwendung gewährleistet ist (Skender 2013).

### **Die Bedeutung von Diagrammen im naturwissenschaftlichen Unterricht und Messung der Diagrammkompetenz**

Mathematische Darstellungen spielen in den Naturwissenschaften eine wichtige Rolle und sind auch in den Kerncurricula der Unterrichtsfächer verankert. Die Veranschaulichung einfacher Beziehungen durch Tabellen und Diagramme sollte bereits im naturwissenschaftlichen Unterricht in den unteren Jahrgangsstufen der Sekundarstufe I verstärkt gefordert und gefördert werden. Diagramme stellen im naturwissenschaftlichen Unterricht eine häufig verwendete Informationsquelle dar. Wenn Lernende sie für Darstellungszwecke selber konstruieren, können sie auch eine lernförderliche Wirkung haben (Stern et al. 2003), müssen es aber nicht (Aprea, Ebner 2003), sodass hier noch weiterer Forschungsbedarf besteht (Lachmayer, Nerdel, Prechtel 2007).

Diagramme repräsentieren Beziehungen zwischen Variablen. In den Naturwissenschaften dienen sie vor allem der Visualisierung quantitativer Zusammenhänge. Dabei ist bei den Achsendiagrammen (Linien-, Punkte-, Balken-, Säulendiagrammen) mindestens eine der Variablen kontinuierlich (Schnotz 1993). Die Beziehungen zwischen den Elementen ergeben sich durch ihre Lage auf den Achsen (Winn 1987). Achsendiagramme besitzen eine durch Konvention festgelegte Form und müssen nach bestimmten Regeln konstruiert und interpretiert werden.

Das Verstehen und die Handhabung von Diagrammen muss entsprechend im Unterricht gefördert werden. Dafür ist es notwendig den Status der bereits vorhandenen Kompetenzen der SchülerInnen zu diagnostizieren.

Bei der in dieser Studie verwendeten Aufgabe müssen die SchülerInnen ihre selbst erhobenen Daten in Form eines Diagramms darstellen. Dies erfordert bestimmte kognitive Fähigkeiten, die sich durch empirische Befunde zur Diagrammkonstruktion charakterisieren lassen. Folgende Schwierigkeiten von SchülerInnen sind bei der Konstruktion von Diagrammen bekannt:

- Das freie Erstellen eines Diagramms aus nur vorgegebenen Daten fällt den SchülerInnen auch in den höheren Jahrgangsstufen noch schwer (Baker, Corbett, Koedinger 2001; Baumert et al. 1999). Die Schwierigkeiten bestehen dabei bereits in der Wahl eines Diagrammtyps, der den Daten angemessen ist.
- Die Konstruktion eines Achsenkreuzes mit entsprechender metrischer Skalierung und das Eintragen der konkreten Datenpunkte bereitet den SchülerInnen Schwierigkeiten bei der Erstellung eines Liniendiagramms (Baker, Corbett, Koedinger 2001; Wavering 1989).
- Problematisch ist darüber hinaus die Unterscheidung, wann Datenpunkte im Sinne einer

Interpolation verbunden werden und wann das Verbinden eher vermieden werden sollte, wenn fälschlicherweise ein Trend bei einer kategorialen Variable angedeutet wird.

Diese Befunde zu den Schwierigkeiten der SchülerInnen beschreiben Anforderungen beim Konstruieren von Diagrammen, die sich zur systematischen Beschreibung von Schülerfähigkeiten modellieren lassen (Lachmayer, Nerdel, Prechtel 2007).

Lachmayer et al. (2007) schlagen ein Strukturmodell der kognitiven Fähigkeiten zum Umgang mit Diagrammen vor, das auf dem integrativen Modell von Schnotz (2001) basiert. Dieses umfasst die drei Bereiche *Informationsentnahme*, *Konstruktion* und *Integration*. Da für die vorliegende Studie im Wesentlichen die eigenständige Konstruktion relevant ist, soll nur dieser Teil des Modells hier detaillierter in seiner Feinstruktur dargestellt werden (s. Tabelle 1).

**Tabelle 1:** Strukturmodell der Diagrammkompetenz für den Bereich der Konstruktion. Durch \* gekennzeichnete Operationen beziehen sich nur auf Liniendiagramme (Lachmayer, Nerdel, Prechtel 2007).

<b>Konstruktion</b>	
Wahl des passenden Diagrammtyps	Aufbau des Rahmens
Zuordnung der Variablen zu ihren Achsen	
Beschriftung der Achsen	
Zeichnen einer Legende	
Zeichnen der Skalen	
Eintragen der Punktwerte	Eintragen der Daten
Skizzierung einer Verbindungslinie zwischen Punkten* oder freie Skizzierung einer Trendlinie	

Die Strukturen hinsichtlich der Konstruktion lassen sich den Fähigkeitsfacetten *Aufbau des Rahmens* und *Eintragen der Daten* differenzierend zuordnen, ohne damit eine Stufung der Fähigkeiten zu postulieren. Die einzelnen Fähigkeiten lassen sich in ein Auswertungsmanual übertragen (s. Tabelle 2 auf S. 9) und dichotom bewerten.

### Untersuchungsdesign

Die experimentell-naturwissenschaftliche Problemlösekompetenz wird anhand eines Realexperiments erhoben. Die SchülerInnen arbeiten dabei in Zweiergruppen und erstellen einzeln, parallel zum Experimentieren prozessorientierte, vorstrukturierte Protokolle. Aus zeitökonomischen Gründen wird für die Messung der kognitiven Fähigkeiten in Anlehnung an andere Studien (Wenzel, 2016; Höner, Wenzel 2018; Emden 2011) ebenfalls eine Kurzform des KFT mit nur je einem Subtest aus dem verbalen, numerischen sowie figural-bildhaften Bereich verwendet (V3, Q2, N3).

Folgende Hypothese wird untersucht:

*Es gibt einen statistisch signifikanten Zusammenhang zwischen der Leistung der schriftlichen Dokumentation beim naturwissenschaftlich-experimentellen Problemlösen und den kognitiven Fähigkeiten im verbalen, numerischen und figural-bildhaften Bereich.*

Im Hinblick auf die Aufgabe zur Diagrammkonstruktion soll folgende Fragestellung untersucht werden:

*Inwiefern verfügen bereits die untersuchten 5. Klässler über die Fähigkeit, aus den erhaltenen Daten ein geeignetes Diagramm zu erstellen, das den Daten angemessen ist?*

Zur Überprüfung der Hypothese werden die schriftlichen Aufzeichnungen der SchülerInnen quantitativ anhand eines bereits vorhandenen Auswertungsmanuals ausgewertet. Der Teil zur Diagrammkonstruktion aus den erhaltenen Daten ist in Anlehnung an das Strukturmodell zur

Diagrammkompetenz für den Bereich Konstruktion angelegt (Lachmayer, Nerdel, Prechtel 2007). Da davon auszugehen ist, dass die schriftlichen Aufzeichnungen der SchülerInnen einer Gruppe recht ähnlich sind, wird das experimentelle Vorgehen zur Erhebung der individuellen naturwissenschaftlich-experimentellen Problemlösekompetenz videografiert und im Anschluss daran qualitativ ausgewertet, sodass verschiedene Perspektiven bezogen auf denselben Untersuchungsgegenstand erhalten werden. Dabei werden die Vorgehensweisen und „Denkprozesse“ der SchülerInnen explorativ beobachtet, um Aspekte herauszuarbeiten, die eventuelle Abweichungen von der schriftlichen Erhebung aufdecken und den Erfolg oder Misserfolg beim experimentellen Problemlösen erklären können (Kelle 2014), sodass sich folgende weitere Forschungsfrage ergibt:

*Welche Faktoren beeinflussen den Erfolg der SchülerInnen beim naturwissenschaftlich-experimentellen Problemlösen?*

### **Stichprobe und Rahmenbedingungen**

An der Untersuchung nahmen 52 SchülerInnen aus zwei Klassen des 5. Jahrgangs einer Integrierten Gesamtschule in Niedersachsen teil. Es handelt sich um eine Gelegenheitsstichprobe. Die beiden Klassen wurden im Fach Naturwissenschaften von derselben Lehrkraft unterrichtet und hatten gerade ihr erstes Schuljahr in diesem Unterrichtsfach hinter sich. Da nur die Ergebnisse der SchülerInnen in die Auswertung einbezogen wurden, die an allen Testteilen teilnahmen, ergibt sich eine Anzahl von insgesamt  $N = 44$ , deren Daten in die Auswertung eingegangen sind. Davon waren 25 Jungen (56,8 %) und 19 Mädchen (43,2 %).

Die teilnehmenden SchülerInnen wurden vor der Durchführung ausdrücklich darauf hingewiesen, dass die Untersuchung keinerlei Einfluss auf die schulische Leistungsbewertung hat und dass ihre Lehrkraft keine Einsicht in die schriftlichen Aufgabenbögen oder die Videodaten erhält.

Für das Realexperiment standen drei Schulstunden á 45 Minuten zur Verfügung, wobei nach zwei Schulstunden jeweils eine Pause von 20 Minuten gemacht wurde. Dieser Teil der Untersuchung fand in den naturwissenschaftlichen Fachräumen der Schule statt. Die Bearbeitung des KFT nahm eine Schulstunde á 45 Minuten in Anspruch und wurde in den Klassenräumen durchgeführt.

### **Testinstrumente**

#### **Kognitiver Fähigkeitstest (KFT)**

In dieser Untersuchung wird der Kognitive Fähigkeitstest *KFT 4-12 +R* von Heller und Perleth (2000) in einer verkürzten Form durchgeführt, in der von den SchülerInnen zu jedem Bereich nur je ein Subtest zu bearbeiten ist (vgl. auch Wenzel 2016; Emden 2011). Im verbalen Bereich wird der Subtest *V3 Wortanalogien*, mit 20 Items, im quantitativen Bereich der Subtest *Q2 Zahlenreihen*, mit ebenfalls 20 Items, und im nonverbalen Bereich der Subtest *N3 Faltaufgaben*, mit 15 Items, für die Untersuchung verwendet, sodass sich zur Bearbeitung der Aufgaben ein Gesamtzeitbedarf von 31:30 Minuten ergibt (Heller, Perleth 2000). Die Instruktionstexte und Beispielaufgaben wurden wie im Manual angegeben besprochen.

#### **Realexperiment zum naturwissenschaftlich-experimentellen Problemlösen**

Mit Hilfe des Realexperiments soll die naturwissenschaftlich-experimentelle Problemlösekompetenz der SchülerInnen des 5. Jahrgangs bestimmt werden. Begleitend zum Experiment sind von den SchülerInnen drei Teilaufgaben schriftlich auf vorgefertigten Arbeitsblättern zu bearbeiten.

Beim Experiment soll der Einfluss der Wassertemperatur auf die Auflösezeit von Brausetabletten untersucht werden. Alle benötigten Materialien sind auf den Schülertischen bzw. auf dem Lehrerpult bereitgestellt. Die Partnerwahl wird den SchülerInnen selbst überlassen, um möglichst zu vermeiden, dass eine mangelnde Teamfähigkeit einzelner Experimentierpaare diese im Vergleich zu anderen Paaren beim Lösen der Aufgaben behindert. Einleitend werden Fragen zu den Materialien geklärt, die

sich überwiegend auf die Funktion und Handhabung des digitalen Thermometers und der Stoppuhr beziehen.

Neben dem Versuchsmaterial findet jede(r) Schüler(in) auf seinem bzw. ihrem Arbeitsplatz einen Umschlag mit vier DIN A4-Seiten vor. Die Umschläge werden gemeinsam geöffnet. Auf der ersten Seite ist lediglich der persönliche Code für die Anonymisierung der Daten zu erstellen und einzutragen, die weiteren drei Seiten sind die Aufgabenblätter zum Experiment. Das erste Aufgabenblatt trägt die Überschrift „Auflösen von Brausetabletten“ sowie folgenden einleitenden Text:

Eine Brausetablette kann in kaltem oder heißem Wasser aufgelöst werden.

***Forschungsfrage: Wie hängt die benötigte Zeit zum Auflösen der Tablette von der Wassertemperatur ab?***

Drei Schülergruppen haben Vermutungen zu dieser Frage aufgestellt:

1. Larissa und Lennart meinen:  
Bei einer höheren Wassertemperatur nimmt die Zeit für das Auflösen zu.
2. Marie und Finn vermuten:  
Bei einer höheren Wassertemperatur nimmt die Zeit für das Auflösen ab.
3. Max und Aila glauben:  
Die Zeit für das Auflösen hat nichts mit der Temperatur zu tun.

Untersuche die Vermutungen experimentell.

Dazu hast du die folgenden Materialien zur Verfügung:

- |                                      |                   |
|--------------------------------------|-------------------|
| ▪ Kaltes Wasser                      | ▪ Messzylinder    |
| ▪ Heißes Wasser aus dem Wasserkocher | ▪ Brausetabletten |
| ▪ Bechergläser                       | ▪ Stoppuhr        |
| ▪ Messbecher (um Wasser zu holen)    | ▪ Thermometer     |

Unter diesem einleitenden Text befindet sich die Aufgabe 1 und darunter ein freies Feld für die Notizen der SchülerInnen mit der Überschrift „Arbeitsschritte“ abgedruckt.

#### **Aufgabe 1:**

- Miss bei mindestens vier unterschiedlichen Temperaturen die Zeit bis zum Auflösen.
- Notiere dabei deine Arbeitsschritte!
- Schreibe deine Messwerte auf Seite 2 in einer geeigneten Form auf.

Auf dem zweiten Aufgabenblatt ist oben ein weiteres zu Aufgabe 1 gehörendes, freies Feld mit der Überschrift „Messwerte“.

Durch den einleitenden Text erhalten die SchülerInnen schon konkrete Anhaltspunkte zur Problemstellung, da die Forschungsfrage und drei verschiedene Vermutungen vorgegeben sind. Diese Vereinfachung wurde gewählt, weil die SchülerInnen der 5. Jahrgangsstufe kaum Erfahrung mit naturwissenschaftlichen Denk- und Arbeitsweisen sowie freiem Experimentieren haben. Darüber hinaus wird der konkrete Hinweis zur Durchführung gegeben, dass die Zeit bis zum Auflösen bei mindestens vier verschiedenen Wassertemperaturen gemessen und die Arbeitsschritte sowie die erhaltenen Messwerte notiert werden sollen.

Mit der zweiten Aufgabe unter dem freien Feld für die Messwerte werden die SchülerInnen aufgefordert, ein Diagramm zu erstellen.

**Aufgabe 2:** Stelle deine Messwerte in einem Diagramm grafisch dar.

Ein Koordinatensystem mit nicht beschrifteter x- und y-Achse ist bereits vorgegeben. Geeignete Darstellungen für die Daten sind Säulen-, Balken-, Punkte- bzw. Liniendiagramme.

Bei der dritten Aufgabe sollen die SchülerInnen einen Rückbezug zu den Vermutungen aus dem Einleitungstext herstellen und auf Grundlage ihrer experimentellen Ergebnisse erklären.

**Aufgabe 3:** Nimm zu den Vermutungen Stellung.

Erkläre aufgrund deiner Ergebnisse den Zusammenhang zwischen der Wassertemperatur und der Zeit für das Auflösen der Brausetablette.

Grundsätzlich sind die drei Teilaufgaben als offene Aufgaben anzusehen, obwohl die Vorgehensweise bereits vorstrukturiert ist.

Das verwendete Realexperiment orientiert sich am SDDS-Modell, wobei die Suche im Hypothesenraum entfällt und mit der Suche im Experimentierraum (Testen von Hypothesen) begonnen wird. Die SchülerInnen müssen einen Versuch planen, mit dem sie die Forschungsfrage beantworten und die Vermutungen überprüfen können. Die dritte Aufgabe ist der Bewertung von Evidenzen zuzuordnen.

Es kann davon ausgegangen werden, dass die Aufgabenstellung für die SchülerInnen ein „Problem“ darstellt, da sie bisher noch nicht forschend im Unterricht experimentiert haben. Eine Barriere auf dem Weg vom Ausgangs- zum Zielzustand ist dabei, dass vier unterschiedliche Wassertemperaturen verwendet werden sollen, obwohl nur heißes und kaltes Wasser zur Verfügung steht, sodass sie durch Mischen vier verschiedene Temperaturen herstellen müssen. Des Weiteren muss bei allen vier Versuchsansätzen das Wasservolumen konstant gehalten werden, um vergleichbare Daten zu erhalten.

Bei der zweiten Aufgabe liegt der Fokus im mathematischen Bereich, da die Daten in einem Diagramm aufbereitet werden sollen. Dieser Aufgabenteil ist für die Beantwortung der Forschungsfrage eigentlich nicht notwendig, da den Messwerten direkt entnommen werden kann, dass sich die Brausetabletten bei einer höheren Temperatur schneller auflösen. Aufgrund der Bedeutung des Aufbaus entsprechender Diagrammkompetenzen im naturwissenschaftlichen Unterricht, erscheint die Einbindung dieser Aufgabe sinnvoll, um Hinweise auf die Diagrammkompetenz zu erhalten (s. Tabelle 2).

### **Quantitative Auswertung der Protokolle und des KFT**

Im Sinne der klassischen Testtheorie sollen die Ergebnisse auf korrelative Zusammenhänge zwischen den Leistungen bei der schriftlichen Dokumentation des naturwissenschaftlich-experimentellen Problemlösens und den kognitiven Fähigkeiten untersucht werden.

Die Protokolle werden entsprechend eines Auswertungsmanuals bewertet. Dieses wurde bereits in anderen Untersuchungen verwendet und hat sich als sehr gut hinsichtlich der Beurteilerübereinstimmung erwiesen (Höner, Matis 2017). Insgesamt sind maximal 18 Punkte zu erreichen, wobei 7 Punkte für die erste Aufgabe, 8 Punkte für die zweite Aufgabe und 3 Punkte für die dritte Aufgabe vergeben werden (s. Tabelle 2).



**Tabelle 2:** Erwartungshorizont für die Punktevergabe bei der naturwissenschaftlich-experimentellen Aufgabensequenz (Höner, Matis 2017).

Erwartungshorizont Aufgabe 1 (je 1 Punkt)	Anmerkungen
<ul style="list-style-type: none"> <li>• Es werden vier verschiedene Wassertemperaturen verwendet.</li> <li>• Die Zeiten zum Auflösen werden gemessen.</li> <li>• Die Daten wurden sinnvoll notiert.</li> <li>• Die Temperaturen und Zeiten zum Auflösen lassen sich eindeutig zuordnen.</li> <li>• Das Vorgehen wurde nachvollziehbar notiert.</li> <li>• Es wurde notiert, dass das Wasservolumen immer gleich war.</li> <li>• Das Experiment bezieht sich auf die Fragestellung.</li> </ul>	<p><i>Unterschiede zwischen den Temperaturen <math>\geq 5\text{ }^{\circ}\text{C}</math>.</i></p> <p><i>Keine schwer nachvollziehbaren Sätzen mit überflüssigen Informationen.</i></p> <p><i>Zeit zum Auflösen wird gemessen, nicht z.B. die Veränderung der Wassertemperatur.</i></p>
Erwartungshorizont Aufgabe 2 (je 1 Punkt)	Anmerkungen
<ul style="list-style-type: none"> <li>• Auf der x-Achse wurde die Temperatur eingetragen.</li> <li>• Auf der y-Achse wurde die Zeit zum Auflösen eingetragen.</li> <li>• Die Achseneinteilung auf der x-Achse ist nach Größe sortiert.</li> <li>• Die Achseneinteilung auf der y-Achse ist nach Größe sortiert.</li> <li>• Auf der x-Achse ist die Einheit angegeben.</li> <li>• Auf der y-Achse ist die Einheit angegeben.</li> <li>• Die Messwertdarstellung ist ansatzweise korrekt.</li> <li>• Die Messwerte wurden richtig in das Diagramm übertragen.</li> </ul>	<p><i>Konvention für die unabhängige Variable. Ausnahme: Balkendiagramme.</i></p> <p><i>Konvention für die abhängige Variable. Ausnahme: Balkendiagramme.</i></p> <p><i>Die Werte sind der Größe nach sortiert. Sie müssen nicht unbedingt metrisch sein.</i></p> <p><i>Die Werte sind der Größe nach sortiert. Sie müssen nicht unbedingt metrisch sein.</i></p> <p><i>Die Größen sind auf der richtigen Achse, aber evtl. ungenau eingetragen.</i></p> <p><i>2 Punkte, da die Kategorie „ansatzweise“ enthalten ist.</i></p>
Erwartungshorizont Aufgabe 3 (je 1 Punkt)	Anmerkungen
<ul style="list-style-type: none"> <li>• Es wird Stellung zu den Vermutungen genommen.</li> <li>• Der Zusammenhang ist richtig im Sinne der Ergebnisse.</li> <li>• Die Stellungnahme wird mit dem Experiment begründet.</li> </ul>	<p><i>Es wird überhaupt Stellung genommen.</i></p> <p><i>Der Zusammenhang zwischen der Wassertemperatur und der Zeit zum Auflösen wird korrekt wiedergegeben.</i></p> <p><i>Das eigene Experiment wird als Begründung herangezogen, z.B. „bei uns war es so...“, „... haben recht, <u>weil</u> sich die Tablette in heißem Wasser schneller auflöst.“</i></p>

Darüber hinaus wurde ebenfalls notiert, ob ein Säulen-, Balken-, Punkte- oder Liniendiagramm erstellt wurde, ohne dafür einen Punkt zu geben, da diese Darstellungen prinzipiell alle geeignet sind. Ebenso wurde ohne Bepunktung geschaut, ob die Abstände auf den Achsen metrisch dargestellt wurden.

Der KFT wird entsprechend der Musterlösungen bewertet. Im verbalen und quantitativen Subtest können jeweils maximal 20 Punkte und im nonverbalen Subtest maximal 15 Punkte erreicht werden. Die maximal zu erreichende Punktzahl beim KFT liegt also bei 55 Punkten.

Für die quantitative Auswertung wurden die Daten in die Statistiksoftware IBM SPSS Statistics 24 (SPSS) eingetragen.

Da die Daten als intervallskaliert angesehen werden können, wurden mit Hilfe deskriptiver Statistik die Mittelwerte und Standardabweichungen berechnet. Aus dem Quotienten des Mittelwertes und der zu erreichenden maximalen Punktzahl wurde für jeden Testteil der Schwierigkeitsindex ermittelt. Bei einem Schwierigkeitsindex von  $< 0,20$  sind Aufgaben als zu schwierig und bei einem Schwierigkeitsindex  $> 0,80$  als zu leicht anzusehen. Ideal sind Aufgaben im mittleren Schwierigkeitsbereich von etwa 0,50 (Maier 2014).

Die vorliegenden Daten wurden mit Hilfe des *Kolmogorov-Smirnov-Anpassungstests* auf Normalverteilung getestet, um zu entscheiden, welcher Korrelationskoeffizient berechnet werden kann. Es werden bivariate, zweiseitige Korrelationen berechnet, die durch den Korrelationskoeffizienten  $r$  sowie durch den p-Wert  $p$  charakterisiert werden. Die Korrelation zwischen zwei Variablen wird für  $p \leq 0,05$  als „schwach signifikant“ und für  $p \leq 0,01$  als „signifikant“ bezeichnet und durch einen Stern (\*) bzw. zwei Sterne (\*\*) gekennzeichnet. Entsprechend der Konvention von Cohen (1988) werden Zusammenhänge mit einem Korrelationskoeffizienten  $r < 0,10$  als unbedeutend angesehen; bei  $r \geq 0,10$  wird von einer geringen, bei  $r \geq 0,30$  von einer mittleren und bei  $r \geq 0,50$  von einer starken Korrelation gesprochen.

### Qualitative Auswertungen der Videodaten

Im Rahmen der qualitativen Auswertung wurden die videografierten Vorgehensweisen einzelner SchülerInnen beim experimentell-naturwissenschaftlichen Problemlösen beobachtet und induktiv ausgewertet.

Es wurden vier Gruppen ausgewählt, deren Punktzahlen aller Untersuchungsteile in Tabelle 3 dargestellt sind. Es wurden je zwei Gruppen ausgewählt, die bei der quantitativen Protokollbewertung entweder besonders gut (Gruppen A und B) oder besonders schlecht (Gruppen C und D) abgeschnitten haben. Die Gruppen wurden so gewählt, dass sowohl bei den starken als auch bei den schwachen Gruppen bei je einer Gruppe die Punktzahlen beider Gruppenmitglieder identisch sind (Gruppen A und C) und bei der jeweils anderen Gruppe unterschiedliche Punktzahlen innerhalb der Gruppe vorliegen (Gruppen B und D).

**Tabelle 3:** In den einzelnen Tests erreichte Punktzahlen der für die qualitative Auswertung ausgewählten SchülerInnen.

	Gruppe A stark, identisch		Gruppe B stark, unterschiedlich		Gruppe C schwach, identisch		Gruppe D schwach, unterschiedlich	
	ANTO10	JEHE10	NIRE07	ANSI05	LUAL09	CAGE10	ANMA12	HIAK06
<b>Aufgaben-sequenz</b> (max. 18 P.)	13	13	13	16	5	5	5	5
<b>Aufgabe 1</b> (max. 7 P.)	7	7	5	6	4	4	1	2
<b>Aufgabe 2</b> (max. 8 P.)	5	5	5	7	0	0	0	0
<b>Aufgabe 3</b> (max. 3 P.)	1	1	3	3	1	1	1	3
<b>KFT</b> (max. 55 P.)	31	44	39	40	18	28	18	42

Die Videoaufnahmen der vier Gruppen wurden zuerst wörtlich transkribiert und in jeder Sequenz beobachtbare Handlungen notiert. Um die Sprachhandlungen und Handlungen der SchülerInnen zu ordnen und zu systematisieren, wurde anschließend ein Kategoriensystem entwickelt.

Aufgrund des relativ großen Umfangs an Datenmaterial sowie wegen der Gefahr, dass durch Verkürzung von Textpassagen der Kontext nicht mehr nachvollziehbar ist, wurde auf die Paraphrasierung inhaltstragender Textstellen verzichtet. Stattdessen wurde der gesamte Text nach für die Forschungsfrage relevanten Passagen durchsucht und die entsprechenden Textstellen markiert. Das Kategoriensystem umfasst 14 Kategorien, die in Tabelle 4 in einer Kurzform (Bezeichnung und Definition) aufgeführt sind.

**Tabelle 4:** Kategorien des Kategoriensystems und deren Definitionen.

Kategorie	Definition
<b>Lehrerinteraktion</b>	Alle Hinweise von Seiten der Lehrkraft, die sich auf den Versuch oder die Aufgaben beziehen.
<b>Dokumentation</b>	Alle Handlungen, bei denen SchülerInnen etwas notieren.
<b>themenfremde Handlungen u. Äußerungen</b>	Äußerungen und Handlungen, die nichts mit dem Unterrichtsinhalt zu tun haben.
<b>zielführende Handlungen</b>	Alle Handlungen, die korrekt ausgeführt werden und sich zielführend hinsichtlich der Aufgabenstellung auswirken.
<b>nicht zielführende Handlungen</b>	Alle Handlungen, die nicht zielführend sind und eine erfolgreiche Versuchsdurchführung behindern könnten.
<b>korrekter Umgang mit Kontrollvariablen</b>	Alle Äußerungen und Handlungen, die auf den korrekten Umgang mit Kontrollvariablen abzielen.
<b>ungenauere Messung</b>	Alle Äußerungen und Handlungen, die auf eine ungenaue Messung oder eine Verfälschung der Messwerte hinweisen.
<b>Zielführende Äußerungen</b>	Alle Äußerungen, die zielführend für die Versuchsdurchführung sind.
<b>nicht zielführende Äußerungen</b>	Alle Äußerungen, die zwar im weitesten Sinne etwas mit dem Versuch zu tun haben, jedoch nicht zielführend sind.
<b>Arbeitsorganisation</b>	Alle Äußerungen, die zielführend für die Arbeitsorganisation sind und ein strukturiertes Vorgehen der Gruppe fördern.
<b>Vermutungen</b>	Alle Vermutungen und deren Begründungen bezüglich des Experiments sowie Äußerungen, die sich auf die im Einleitungstext auf dem ersten Aufgabenblatt genannten Vermutungen beziehen.
<b>Diagramm</b>	Alle Äußerungen, die zielführend für die graphische Darstellung der Messwerte sind.
<b>mangelnde Selbstständigkeit</b>	Alle Äußerungen und Handlungen, die auf mangelnde Selbstständigkeit oder mangelndes selbstständiges Denken hinweisen.
<b>Verständnis-schwierigkeiten</b>	Alle Äußerungen, die auf inhaltliche Verständnisschwierigkeiten hinweisen.

Auf Grundlage der vorgenommenen Kategorisierung wurden die relevanten Textpassagen im Rahmen einer qualitativen Inhaltsanalyse interpretiert. Die Ergebnisse der einzelnen Gruppen wurden explorativ miteinander verglichen, um mögliche Gemeinsamkeiten, Unterschiede oder Auffälligkeiten zu ermitteln und die Ergebnisse der quantitativen Untersuchung zu ergänzen.

## Ergebnisse der quantitativen Auswertung

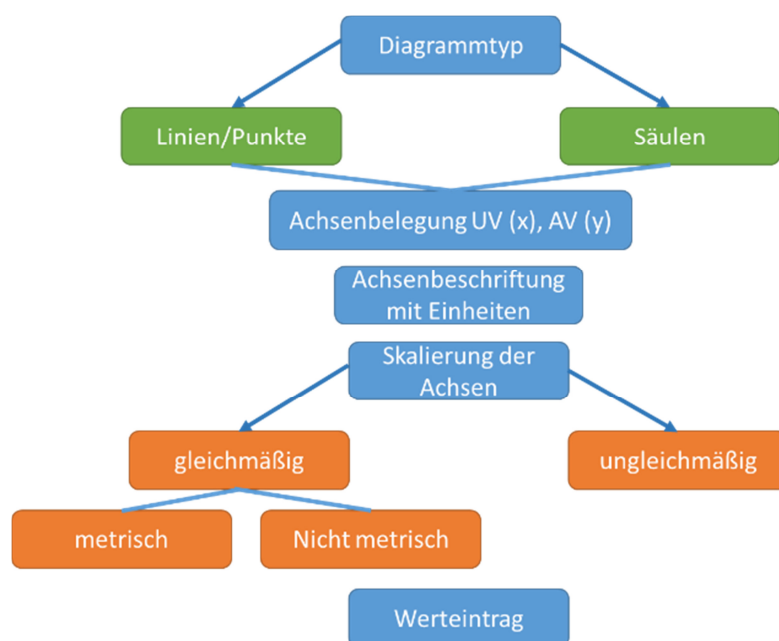
Die deskriptive Auswertung ergibt folgende Mittelwerte, Standardabweichungen und Schwierigkeitsindizes für die (Sub-)Tests bzw. Aufgaben (s. Tabelle 5).

**Tabelle 5:** Mittelwerte, Standardabweichungen, maximal erreichbare Punkte und Schwierigkeitsindizes der einzelnen Tests (N = 44).

	Mittelwert ± Standardabweichung	Max. Punkte	Schwierigkeitsindex
verbaler Subtest	10,11 ± 4,16	20	0,506
quantitativer Subtest	16,27 ± 3,11	20	0,814
nonverbaler Subtest	9,00 ± 4,34	15	0,600
KFT	35,39 ± 8,73	55	0,643
Aufgabe 1	4,70 ± 1,19	7	0,671
Aufgabe 2	2,73 ± 2,20	8	0,341
Aufgabe 3	1,52 ± 1,07	3	0,507
Aufgabensequenz	8,95 ± 3,18	18	0,497

Die Schwierigkeitsindizes der Subtests des KFT zeigen, dass sowohl der verbale als auch der nonverbale Subtest eine gute mittlere Schwierigkeit aufweisen. Der quantitative Subtest war dagegen mit 0,814 zu leicht. Insgesamt besitzt der KFT einen Schwierigkeitsindex von 0,643. Die Berechnungen der Schwierigkeitsindizes der experimentellen Aufgaben ergeben, dass die erste Aufgabe etwas zu leicht und die zweite Aufgabe etwas zu schwierig war, während die dritte Aufgabe im mittleren Schwierigkeitsbereich lag. Bezogen auf die gesamte Aufgabensequenz wurde ein Schwierigkeitsindex von 0,497 errechnet, sodass die experimentelle Aufgabensequenz insgesamt eine gute mittlere Schwierigkeit aufweist.

Die Aufgabe 2 zur Diagrammkonstruktion wurde entsprechend eines Entscheidungsbaums (s. Abbildung 2) in Anlehnung an das Strukturmodell (s. Tabelle 1) auch nach prozentualen Häufigkeiten untersucht. Da in keinem Fall ein Balkendiagramm erstellt wurde, bleiben diese hier unberücksichtigt.



**Abbildung 2:** Entscheidungsbaum (in Anlehnung an Lachmayer, Nerdel, Prechtel 2007)

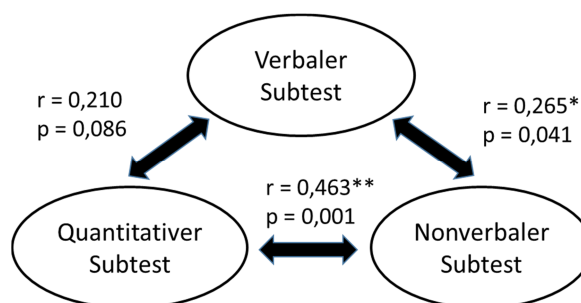
Tabellarisch ergibt sich folgender prozentualer Anteil der SchülerInnen (N = 44) bei der Feingliederung der Diagrammkonstruktion. 22,2 % der SchülerInnen erstellten kein Diagramm.

**Tabelle 6:** Prozentale Anteile verschiedener Diagrammteile.

Aspekt	% SchülerInnen (N = 44)
Säulendiagramm	42,2
Linien/Punktediagramm	35,6
Achsenbelegung: Unabhängige Variable UV (Temperatur) auf x-Achse	40,0
Achsenbelegung: Abhängige Variable AV (Zeit) auf y-Achse	40,0
Achsenbeschriftung mit Einheit angegeben: x-Achse	35,6
Achsenbeschriftung mit Einheit angegeben: y-Achse	37,8
Sinnvolle (metrische) x-Achseneinteilung	15,6
Sinnvolle (metrische) y-Achseneinteilung	48,9
Metrische Einteilung der x-Achse bei Linien/Punktediagramm	35,6
Metrische Einteilung der y-Achse bei Linien/Punktediagramm	53,3
Werteeintrag ansatzweise richtig	33,3
Werteeintrag korrekt	17,8

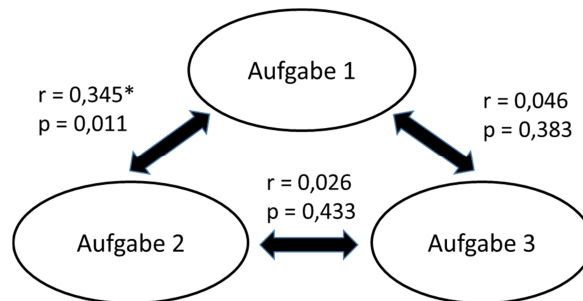
Als Diagrammtyp wurde häufiger das Säulendiagramm gewählt als ein Punkte- bzw. Liniendiagramm, wobei der Unterschied nicht signifikant ist. Bei den Säulendiagrammen wurden von 13,6 % der SchülerInnen jeweils zwei Diagramme erstellt, indem es vier Säulen für die Temperatur gab und vier weitere Säulen für die Auflösezeit. 40 % der SchülerInnen trugen die unabhängige Variable auf der x-Achse ein und entsprechend die abhängige Variable auf der y-Achse. Auffällig ist, dass unabhängig von der Achsenbelegung (abhängige, unabhängige Variable) die metrische Einteilung der y-Achse in 48,9 % bzw. 53,3 % der Fälle richtig erfolgte, während es auf der x-Achse nur 15,6 % bzw. 35,6 % sind. Dies liegt daran, dass bei den Säulendiagrammen häufig einfach die Reihenfolge der Messdaten bei den verschiedenen Temperaturen übernommen wurde, ohne diese in Reihenfolge zunehmender oder abnehmender Temperatur zu sortieren. Einem Drittel der SchülerInnen gelang eine ansatzweise richtige Messwertdarstellung, korrekt war der Dateneintrag nur in 17,8 % der Fälle.

Eine Untersuchung der korrelativen Zusammenhänge zwischen den erreichten Punktzahlen der drei kognitiven Subtests des KFT ergab, dass ein signifikanter Zusammenhang zwischen dem quantitativen und dem nonverbalen Subtest ( $r = 0,463^{**}$  mit  $p = 0,001$ ) existiert mit einem mittleren Korrelationskoeffizienten. Der verbale Subtest korreliert schwach mit dem nonverbalen, aber nicht mit dem quantitativen Subtest. Aus diesem Grund werden in der folgenden Auswertung jeweils nicht nur Zusammenhänge mit dem gesamten KFT, sondern auch mit den einzelnen Subtests überprüft.



**Abbildung 3:** Korrelationen zwischen den kognitiven Aufgabenteilen.

Bei der Überprüfung der Zusammenhänge zwischen den erreichten Punkten in den drei Aufgaben zum Experimentieren konnte ein signifikanter Zusammenhang zwischen Aufgabe 1 und Aufgabe 2 ( $r = 0,345^*$  mit  $p = 0,011$ ) festgestellt werden. Die dritte Aufgabe korreliert mit keiner der anderen Aufgaben, weshalb in der folgenden Auswertung die experimentellen Aufgaben jeweils auch einzeln betrachtet werden.



**Abbildung 4:** Korrelationen zwischen den experimentellen Aufgabenteilen.

Die Korrelationsanalyse der erreichten Punkte in den kognitiven und experimentellen Aufgaben (Tabelle 7) zeigt, dass es einen schwach signifikanten Zusammenhang zwischen dem verbalen Subtest und den in der dritten Aufgabe erreichten Punkten gibt ( $r = 0,309^*$  mit  $p = 0,02$ ). Zudem korreliert der quantitative Subtest mit einem hohen Korrelationskoeffizienten signifikant mit der Punktzahl der dritten Aufgabe ( $r = 0,513^{**}$  mit  $p < 0,001$ ). Der nonverbale Subtest weist einen signifikanten Zusammenhang mit der ersten Aufgabe ( $r = 0,395^{**}$  mit  $p = 0,004$ ) sowie einen schwach signifikanten Zusammenhang mit der gesamten Aufgabensequenz ( $r = 0,303^*$  mit  $p = 0,023$ ) auf.

**Tabelle 7:** Signifikante Korrelationen zwischen den kognitiven und experimentellen Aufgaben ( $N = 44$ ).

			Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgaben-sequenz
Spearman-Rho	Verbaler Subtest	Korrelationskoeffizient Sig. (2-seitig)			0,309* 0,020	0,259* 0,044
	Quantitativer Subtest	Korrelationskoeffizient Sig. (2-seitig)			0,513** 0,000	0,266* 0,041
	Nonverbaler Subtest	Korrelationskoeffizient Sig. (2-seitig)	0,395** 0,004			0,303* 0,023
	KFT	Korrelationskoeffizient Sig. (2-seitig)	0,329** 0,015		0,379** 0,006	0,340* 0,012

Insgesamt zeigt sich damit, dass SchülerInnen, die im KFT eine hohe (niedrige) Punktzahl erreicht haben, signifikant häufiger gut (schlecht) in der gesamten experimentellen Aufgabensequenz abschneiden. Der Korrelationskoeffizient liegt von der Höhe im mittleren Bereich ( $r = 0,340^*$ ).

### Ergebnisse der qualitativen Auswertung

Basierend auf dem in Tabelle 4 dargestellten Kategoriensystem erfolgte eine im Hinblick auf die Fragestellung gezielte Auswertung. Um einen Vergleich zwischen den Gruppen zu ermöglichen, wurde zunächst für jede Gruppe einzeln deren vollständiges Transkript hinsichtlich der Häufigkeiten auftretender Kategorien analysiert. Die so erhaltenen Ergebnisse wurden miteinander verglichen, um mögliche Unterschiede zwischen Gruppen mit hoher und niedriger Punktzahl zu identifizieren. Dabei ergaben sich folgende Aspekte als mögliche relevante Einflussgrößen für die Erfolgswahrscheinlichkeit beim experimentellen Problemlösen: (*Zeitpunkt der ersten*) *Dokumentation*, *Arbeitsorganisation*, *Verständnisschwierigkeiten* und *Diagramm*. In allen vier

Kategorien ließen sich Unterschiede zwischen den Gruppen beobachten. Im Vergleich zu den schwachen Gruppen begannen die starken Gruppen schon frühzeitig mit der *Dokumentation* und sie hatten darüber hinaus einen höheren *Grad an Arbeitsorganisation* und machten deutlich häufiger zielführende Äußerungen bezüglich der zweiten Aufgabe (Kategorie „Diagramm“). Die beiden schwachen Gruppen hatten bezüglich aller beobachteten Aufgabenteile *Verständnisschwierigkeiten*, jedoch hatten auch die starken Gruppen Probleme, die inhaltlichen Zusammenhänge zu verstehen. Diese Schwierigkeiten lagen bei einer starken Gruppe im Bereich der zweiten Aufgabe und bei der anderen im Bereich der dritten Aufgabe. Die bessere Arbeitsorganisation zeigte sich z. B. darin, dass bereits während des Lehrervortrags überlegt wurde, wofür sie das Thermometer und die Stoppuhr benutzen könnten. Weiterhin machten sich diese SchülerInnen darüber Gedanken, in welchem Verhältnis sie sinnvoll das kalte und warme Wasser mischen könnten, um letztendlich vier verschiedene Wassertemperaturen zu erhalten und auch die Variable Wasservolumen wurde meist beachtet.

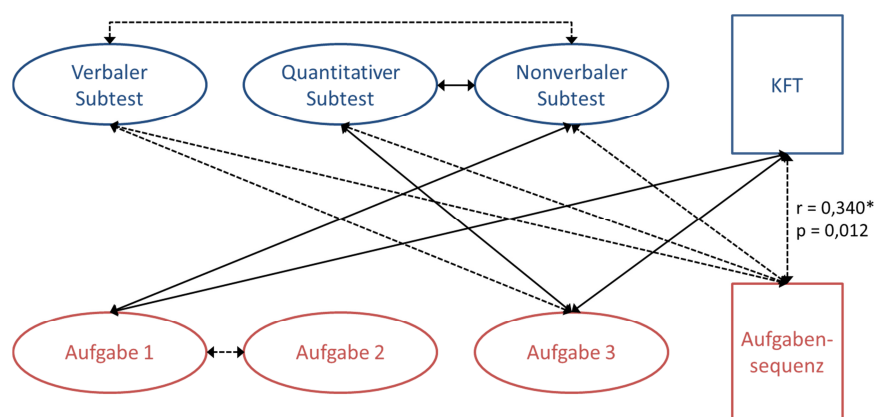
Da bei der quantitativen Auswertung aufgefallen ist, dass die beiden SchülerInnen einer Gruppe häufig eine ähnliche oder dieselbe Punktzahl erreicht haben, wurden bei der qualitativen Auswertung die beiden ExperimentierpartnerInnen einer Gruppe auch einzeln betrachtet. Hierzu wurden die als relevant beurteilten Textpassagen den einzelnen SchülerInnen der jeweiligen Gruppe zugeordnet, gegenübergestellt und schließlich mit den mit Hilfe der SchülerInnen-Protokolle ermittelten Ergebnissen verglichen.

Die Analyse ergab für die Gruppen mit identischen Punktzahlen beider ExperimentierpartnerInnen (Gruppen A und C), dass sich zwar Unterschiede in den einzelnen Experimentierphasen zeigen, dass diese aber insgesamt ausgeglichen sind. D. h. die beiden PartnerInnen ergänzen sich jeweils in ihren Schwächen und Stärken, sodass durch die Punkte nicht wiedergespiegelt wird, bei welchen Aufgabenteilen jeweils individuelle Defizite bestehen. Insgesamt kann aber über die ganze Aufgabe derselbe Leistungsstand der jeweiligen ExperimentierpartnerInnen bestätigt werden.

Ein ähnliches Bild ergab sich für die beiden Gruppen mit unterschiedlicher Punktzahl der beiden ExperimentierpartnerInnen (Gruppen B und D). Auch hier zeigte die qualitative Analyse, dass der Partner mit der jeweils höheren Punktzahl leistungsstärker bei der Aufgabenbearbeitung war. D. h. auch hier wird zwar durch die Bepunktung der Dokumentation kein individuelles, detailliertes Bild der individuellen Leistungsfähigkeit aufgezeigt, aber es findet eine Differenzierung in die richtige Richtung statt.

### Zusammenfassung und Interpretation der Ergebnisse

Bei der quantitativen Auswertung wurden einige signifikante bzw. schwach signifikante Zusammenhänge zwischen den Subtests und einzelnen Aufgaben festgestellt, die in der Abbildung 5 zusammenfassend dargestellt sind. Die Korrelationskoeffizienten liegen dabei von der Höhe her überwiegend im mittleren Bereich.



**Abbildung 5:** Korrelationen zwischen kognitiven und experimentellen Aufgaben(teilen). Durchgezogene Pfeile stehen für signifikante ( $p < 0,01$ ) und gestrichelte Pfeile für schwach signifikante Korrelationen ( $p < 0,05$ ).

Die gesamte Aufgabensequenz weist eine schwach signifikante Korrelation von  $p = 0,012$  mit  $r = 0,340^*$  mit dem gesamten KFT auf. Die erste Hypothese, dass es einen statistisch signifikanten Zusammenhang zwischen der naturwissenschaftlich-experimentellen Problemlösekompetenz und den kognitiven Fähigkeiten im verbalen, numerischen und figural-bildhaften Bereich gibt, kann also durch die Ergebnisse dieser Untersuchung, so wie auch schon in anderen Studien (Wenzel 2016; Höner, Wenzel 2018; Höner et al. 2018), bestätigt werden. Andere mögliche Einflussfaktoren wie z. B. die Motivation wurden hier nicht mit untersucht.

Interessanterweise weist die Aufgabe 2 *Diagrammkonstruktion* nur eine signifikante Korrelation mit Aufgabe 1 auf und gar keine signifikanten Zusammenhänge mit den Subtests des KFT und auch nicht mit dem gesamten KFT. Dies könnte darauf hindeuten, dass für die Diagrammkompetenz andere kognitive Fähigkeiten benötigt werden, die durch die verwendeten Skalen des KFT nicht getestet werden. Die signifikante Korrelation zur Aufgabe 1 zeigt, dass eine sorgfältige Auseinandersetzung mit Planung und Dokumentation des Experiments zu einer leichteren Bewältigung der Diagrammkonstruktion führt.

Die Schwierigkeitsanalyse hat gezeigt, dass es sich bei der Diagrammkonstruktion um den schwierigsten Aufgabenteil handelt. In der Gesamtheit liegt die Schwierigkeit der Problemlöseaufgabe im guten mittleren Bereich, sodass es sinnvoll erscheint, die Aufgabe 2 als differenzierendes Element zwischen potentiell leistungsstarken und eher leistungsschwachen SchülerInnen beizubehalten.

Die Ergebnisse aus Aufgabe 3 korrelieren stark ( $r = 0,513^{**}$ ) mit den Punkten des quantitativen Subtests. Ein Grund dafür könnte sein, dass es um das schlussfolgernde, logische Denken geht, das bei beiden Testteilen eine Rolle spielt. Dabei ist allerdings zu beachten, dass der quantitative Subtest für die Probanden etwas zu leicht war, wodurch es bei den SchülerInnen mit hoher Punktzahl zu Deckeneffekten gekommen sein könnte. Die beiden weiteren Subtests des KFT, die einzelnen Aufgaben und die gesamte Aufgabensequenz besaßen hingegen einen für die SchülerInnen angemessenen Schwierigkeitsgrad.

Bei der qualitativen Auswertung wurden vorrangig Antworten auf die Frage gesucht, welche Faktoren den Erfolg der SchülerInnen beim Lösen der Aufgaben zur naturwissenschaftlich-experimentellen Problemlösekompetenz beeinflussen. Dazu wurde zunächst ein Kategoriensystem entworfen, in dessen Kategorien jegliche Äußerungen und Handlungen der vier untersuchten Gruppen eingeordnet wurden, um so eventuelle Unterschiede zwischen schwachen und starken Gruppen zu identifizieren. Zudem wurden jeweils die relevanten Handlungen und Äußerungen der beiden ExperimentierpartnerInnen miteinander verglichen und diese Ergebnisse mit den in den einzelnen Aufgaben erreichten Punktzahlen abgeglichen.

Anhand der Beobachtungen können einige Hypothesen abgeleitet werden, die erklären, welche Faktoren den Erfolg der SchülerInnen beim naturwissenschaftlich-experimentellen Problemlösen beeinflussen könnten:

- Ein hoher Grad an Arbeitsorganisation wirkt sich positiv auf die Erfolgswahrscheinlichkeit aus.
- Eine frühzeitige Dokumentation des Vorgehens wirkt sich positiv auf die Erfolgswahrscheinlichkeit aus.
- Vermehrte Verständnisschwierigkeiten wirken sich negativ auf die Erfolgswahrscheinlichkeit aus.
- Zeigen beide ExperimentierpartnerInnen unterschiedliche Leistungen, werden diese mit Hilfe der SchülerInnen-Protokolle nicht detailliert erfasst, sondern nur im Rahmen der Gesamtbeurteilung.

Da alle vier beobachteten Gruppen im Verlauf der Videos Verständnisschwierigkeiten gezeigt haben, ist nicht davon auszugehen, dass diese grundsätzlich die Erfolgswahrscheinlichkeit beeinträchtigen.



Dies entspricht auch dem Konzept des Problemlösens, demzufolge ein Problem nur dann vorliegt, wenn dieses nicht durch Routinehandlungen gelöst werden kann. Entsprechend wirkt es sich nicht generell erfolgsmindernd aus, wenn SchülerInnen beispielsweise zunächst Schwierigkeiten haben, die Aufgabenstellung zu verstehen oder die Messwerte graphisch darzustellen, solange sie im Laufe des Problemlöseprozesses eine Strategie entwickeln. Nichtsdestotrotz liegt die Vermutung nahe, dass sich gehäufte Verständnisschwierigkeiten, insbesondere, wenn sie sich durch die gesamte Bearbeitungsphase zu einer Aufgabe hindurchziehen, negativ auf die Erfolgswahrscheinlichkeit auswirken.

## Diskussion und Fazit

In dieser Arbeit wurde untersucht, inwieweit es bei SchülerInnen der 5. Klasse einer niedersächsischen IGS einen statistisch signifikanten Zusammenhang zwischen der Leistung in der schriftlichen Dokumentation beim naturwissenschaftlich-experimentellen Problemlösen und den kognitiven Fähigkeiten im verbalen, numerischen und figural-bildhaften Bereich gibt. Darüber hinaus wurden die Fähigkeiten hinsichtlich des Erstellens eines Diagramms untersucht. Die naturwissenschaftlich-experimentelle Problemlösekompetenz wurde anhand eines bereits erprobten Realexperiments mit SchülerInnen-Protokollen sowie mit Videografiedaten beim realen Experimentieren erhoben. Die quantitativen Berechnungen dienten dazu korrelative Zusammenhänge aufzudecken. Mit der qualitativen Analyse wurde darüber hinaus der Frage nachgegangen, welche Faktoren die Erfolgswahrscheinlichkeit der SchülerInnen beim experimentellen Problemlösen beeinflussen.

Bei den Korrelationsanalysen wurde ein positiver, statistisch signifikanter Zusammenhang mit einem Korrelationskoeffizienten mittlerer Höhe zwischen den Leistungen bei der schriftlichen Dokumentation beim naturwissenschaftlich-experimentellen Problemlösen und den kognitiven Fähigkeiten festgestellt. Dieses Ergebnis stützt die Arbeiten von Höner und Wenzel (2018) und entspricht darüber hinaus den Erkenntnissen von Kretzschmar et al. (2014), denen zufolge sowohl das Leseverständnis als auch die komplexe Problemlösekompetenz durch die kognitiven Fähigkeiten von SchülerInnen vorhergesagt werden kann (Kretzschmar, Neubert, Greiff 2014). Wirth et al. (2005) kommen ebenfalls zu einem ähnlichen Schluss und meinen, dass die kognitiven Fähigkeiten („Intelligenz“) in einem signifikanten Zusammenhang mit dem Problemlösen stehen, wobei sie ihre These durch die sogenannte *Elshout-Raaheim-Hypothese* eingrenzt, nach der die Korrelation zwischen Intelligenz und Problemlösen dann am größten ist, wenn der Problemlösende weder zu wenig, noch zu viel Wissen besitzt (Klieme, Hartig, Wirth 2005). Auch Scherer (2014 a, b) hat die kognitiven Kompetenzen im figural-bildhaften Bereich („figurales Reasoning“) mit einem Subtest des KFT nach Heller und Perleth (2000) erfasst und Zusammenhänge mit der komplexen Problemlösekompetenz festgestellt. Weiterhin zeigen sich für experimentelle Problemlösefähigkeiten in neueren Studien positive statistische Zusammenhänge zu anschauungsgebundenem Denken (Scherer 2014a, 2014b) und zu arithmetischen Grundkompetenzen (Kretzschmar, Neubert, Greiff 2014). Heitmann (2012) identifizierte auch sprachliche Fähigkeiten als starken Prädiktor für naturwissenschaftliche Problemlösefähigkeiten. Dieses Ergebnis stimmt zum Teil mit den beschriebenen kognitiven Teilkompetenzen von Prenzel et al. (2001) überein, da dort unter anderem beschrieben wird, dass sprachliche Fähigkeiten beim „naturwissenschaftlichen Denken, Verstehen und Schlussfolgern“ (Prenzel et al. 2001, 209) von großer Bedeutung sind.

Die in dieser Arbeit gefundenen Zusammenhänge passen zu denen anderer Studien zu diesem Themenfeld, sodass die kognitiven Fähigkeiten als relevante Einflussgröße für die experimentelle Problemlösekompetenz anzusehen sind. Hier ist zu beachten, dass der quantitative Subtest des KFT für die SchülerInnen am Ende der fünften Jahrgangsstufe zu leicht war, weshalb es hier möglicherweise zu Deckeneffekten gekommen sein könnte. Für zukünftige Untersuchungen scheint es beim quantitativen Subtest deshalb ratsam, den für die nächsthöhere Jahrgangsstufe vorgesehenen Test zu verwenden.

Bei der Diagrammkompetenz im Hinblick auf die Konstruktion zeigte sich, dass gut ein Fünftel der 5. Klässler nicht in der Lage war, ein Diagramm ansatzweise zu erstellen. Die Videografiedaten belegen, dass dieses nicht mit mangelnder Motivation zusammenhängt, sondern dass es Verständnisprobleme gab. Die SchülerInnen konnten mit den Begriffen „grafisch“ und „Diagramm“ nichts anfangen, obwohl diese im Mathematikunterricht der 5. Jahrgangsstufe behandelt wurden. Von den drei Aufgabenstellungen war die Diagrammkonstruktion der schwierigste Teil, wie der Schwierigkeitsindex zeigt. Dieses passt zu den Ergebnissen anderer Studien, wonach das freie Erstellen eines Diagramms aus vorliegenden Daten für SchülerInnen schwierig ist (Baker, Corbett, Koedinger 2001; Baumert et al. 1999). In Übereinstimmung mit der Literatur wurde ebenfalls festgestellt, dass die Achsenbelegung eher schwierig ist (Padilla et al. 1986), ebenso die metrische Skalierung und das Eintragen der konkreten Datenpunkte (Baker et al. 2001; Wavering 1989). In Anlehnung an Lachmayer et al. (2007) zeigte sich, dass die Strukturen der Diagrammkompetenz für Linien/Punkte- und Säulendiagramme gleichermaßen gelten. D. h. es gibt keine bedeutsamen Unterschiede bei der Schwierigkeit der Konstruktion, sodass der Diagrammtyp kein zu berücksichtigender Einflussfaktor ist. Anders als bei Lachmayer et al. (2007), die als einen Prädiktor für die Diagrammkompetenz die *Rechenfertigkeit* feststellten, zeigte sich hier kein statistisch signifikanter Zusammenhang zum quantitativen Teil des KFT, der wohl in dieser Studie am ehesten noch dieses Konstrukt darstellt.

Insgesamt weist die in dieser Studie verwendete naturwissenschaftlich-experimentelle Problemlöseaufgabe eine gute mittlere Schwierigkeit auf und ist für den Einsatz in einer 5. Jahrgangsstufe gut geeignet. Der Aufgabenteil zur Diagrammkonstruktion eignet sich gut, um ein Bild von den Fähigkeiten der SchülerInnen in diesem Bereich zu erhalten.

Die qualitative Untersuchung der Forschungsfrage, welche Faktoren die Erfolgswahrscheinlichkeit der SchülerInnen beim naturwissenschaftlich-experimentellen Problemlösen beeinflussen, hat einige Aspekte hervorgebracht, die bei der quantitativen Auswertung nicht so detailliert zu erfassen sind. Bezüglich der relevanten Einflussfaktoren konnten zunächst die Hypothesen abgeleitet werden, dass sich ein hohes Maß an *Arbeitsorganisation* sowie eine *frühzeitige Dokumentation* des Vorgehens positiv und vermehrte *Verständnisschwierigkeiten* negativ auf die Erfolgswahrscheinlichkeit auswirken. Darüber hinaus lassen die qualitativen Beobachtungen allerdings vermuten, dass Leistungsunterschiede zwischen zwei ExperimentierpartnerInnen in den verschiedenen Phasen durch die Protokolle nur unvollständig erfasst werden.

Grundsätzlich muss bei der Einschätzung der hier genannten Befunde berücksichtigt werden, dass es sich um eine Gelegenheitsstichprobe von geringem Umfang handelt, sodass diese nicht generell auf andere SchülerInnen übertragen werden können.

Für die quantitative Auswertung der SchülerInnen-Protokolle ist es erstrebenswert, dass diese in Zukunft die Leistungen der beiden SchülerInnen, die miteinander arbeiten, differenzierter abbildet. Dazu müsste garantiert werden, dass die Durchführung in Einzelarbeit erfolgt.

## Literatur

- Aprea, C. & Ebner, H. G. (2003): *Generierung von Diagrammen als Lernhandlung: Effekt eines Kurzzeittrainings zur Förderung der Text-Graphik-Transformation*. In F. Achtenhagen & E. G. John (Eds.), *Meilensteine der beruflichen Bildung (Vol 1)* (pp. 117-138). Bielefeld: Bertelsmann.
- Baker, R. S., Corbett, A. T. & Koedinger, K. R. (2001): *Toward a model of learning data representations*. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 45-50). Mahwah, NJ: Erlbaum
- Baumert, J., Bos, W., Klieme, E., Lehmann, R., Lehrke, M., Hosenfeld, I., et al. (Eds.). (1999). *Testaufgaben zu TIMSS/III Mathematisch-naturwissenschaftliche Grundbildung und voruniversitäre Mathematik und Physik der Abschlußklassen der Sekundarstufe II (Population 3)*. Berlin: Max-Planck-Institut für Bildungsforschung

- Baur, A. (2015): *Inwieweit eignen sich bisherige Diagnoseverfahren des Bereichs Experimentieren für die Schulpraxis?* Erschienen in: ZDB: Biologie Lehren und Lernen – Zeitschrift für Didaktik der Biologie, Band 19, Nr. 1 (2015), Bielefeld: Universität Bielefeld, S. 30 – 35.
- Becker-Carus, C. & Wendt, M. (2017): *Allgemeine Psychologie. Eine Einführung*. 2. Auflage 2017. Berlin: Springer-Verlag. S. 462 – 466.
- Betsch, T., Funke, J.; & Plessner, H. (2011): *Denken – Urteilen, Entscheiden, Problemlösen. Allgemeine Psychologie für Bachelor*. Berlin, Heidelberg: Springer-Verlag, S. 137 – 150.
- Cohen, J. (1988): *Statistical Power Analysis for the Behavioral Sciences*. 2. Auflage, Hillsdale, New Jersey: Lawrence Erlbaum Associates, S. 82.
- Dörner, D. (1976): *Problemlösen als Informationsverarbeitung*. 1. Auflage, Stuttgart, Berlin, Köln, Mainz: Kohlhammer Verlag, S. 10.
- Emden, M. (2011): *Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*. Berlin: Logos (118).
- Greiff, S. (2012): *Individualdiagnostik komplexer Problemlösefähigkeit*. Waxmann Verlag
- Greiff, S. & Fischer, A. (2013): *Der Nutzen einer komplexen Problemlösekompetenz: Theoretische Überlegungen und empirische Befunde*. Zeitschrift für Pädagogische Psychologie, 27 (1 - 2), S. 27 - 39.
- Hammann, M. (2007): *Das Scientific Discovery as Dual Search-Modell*. Erschienen in: D. Krüger und H. Vogt (Herausgeber): *Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden*. Heidelberg, Berlin: Springer-Verlag, S. 187 – 196.
- Heitmann, P. (2012): *Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*. Dissertation. Humboldt-Universität zu Berlin.
- Heller, K. A. & Perleth, C. (2000): *KFT 4-12+R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision, Manual*. 3. Auflage, Göttingen: Beltz Test, S. 8f.
- Höner, K.; Matis, K. (2017): *Auswertungsmanual für Logicals* (unveröffentlicht), Institut für Fachdidaktik der Naturwissenschaften, Abt. Chemie und Chemiedidaktik, TU Braunschweig.
- Höner, K.; Matis, K. (2017): *Auswertungsmanual für prozessbezogene Schülerprotokolle zur Temperaturabhängigkeit der Lösungsgeschwindigkeit von Brausetabletten* (unveröffentlicht), Institut für Fachdidaktik der Naturwissenschaften, Abt. Chemie und Chemiedidaktik, TU Braunschweig.
- Höner, K. & Wenzel, D. (2018): *Messung naturwissenschaftlich-experimenteller Problemlösefähigkeit und deren Zusammenhang mit kognitiven und nicht-kognitiven Persönlichkeitsmerkmalen von Schülerinnen und Schülern*. ZFdN eingereicht
- Höner, K., Eghtessad, A., Hilfert-Rüppell, D. & Kraeva, L. (2018): *Naturwissenschaftliches Potenzial? – Diagnose von Schülerfähigkeiten zum experimentellen Problemlösen*. Journal für Begabtenförderung, im Druck
- Kelle, U. (2014): *Mixed Methods*. In: N. Baur, J. Blasius (Hg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS Verlag für Sozialwissenschaften, S. 153-166.
- Kipman, U., Kohlböck, G. & Weilguny, W. (2012). *Psychologische Testverfahren zur Messung intellektueller Begabung*. Salzburg: Österreichisches Zentrum für Begabtenförderung und Begabungsforschung (ÖZBF), S. 111ff.
- Klahr, D. (2000): *Exploring Science: The Cognition and Development of Discovery Processes*. Cambridge, MA: MIT Press.
- Klahr, D. & Dunbar, K. (1988): *Dual space search during scientific reasoning*. Cognitive Science 12: S. 1–48

- Klieme, E., Hartig, J. & Wirth, J. (2005): Problemlösekompetenz von SchülerInnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-200-Studie. 1. Auflage 2005, Wiesbaden: Springer VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH, S. 7 – 31.
- Kretzschmar, A., Neubert, J. C. & Greiff, S. (2014): *Komplexes Problemlösen, schulfachliche Kompetenzen und ihre Relation zu Schulnoten*. Zeitschrift für Pädagogische Psychologie, 28 (4), S. 205–215.
- Lachmayer, S. (2008): *Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht*. Dissertation, Christian-Albrechts-Universität zu Kiel.
- Lachmayer, S., Nerdel, C. & Prechtel, H. (2007): *Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht*. ZfDN, Jg. 13, S. 145-160.
- Leutner, D., Klieme, E., Meyer, K. & Wirth, J. (2005): *Die Problemlösekompetenz in den Ländern der Bundesrepublik Deutschland*. In: PISA-Konsortium Deutschland (Hrsg.), PISA 2003: Der zweite Vergleich der Länder in Deutschland - Was wissen und können Jugendliche? Münster: Waxmann, S. 125-146.
- Leutner, D., Fleischer, J., Wirth, J., Greiff, S. & Funke, J. (2012): *Analytisches und dynamisches Problemlösen im Lichte internationaler Schulleistungsvergleichsstudien*. Psychologische Rundschau, 63, S. 34-42.
- Lotz, M. (2016): *Beschreibung und Beurteilung von Unterricht durch Videoanalysen*. In: Kognitive Aktivierung im Leseunterricht der Grundschule, Wiesbaden: Springer VS Verlag für Sozialwissenschaften, S. 149 – 159.
- Maier, U. (2014): *Leistungsdiagnostik in Schule und Unterricht*. Reihe: Studententexte Bildungswissenschaft, Verlag Julius Klinkhardt: München, S. 54.
- Mayer, J., Grube, C. & Möller, A. (2008): *Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung*. In: U. Harms, A. Sandmann (Hg.), Lehr- und Lernforschung in der Biologiedidaktik Bd.3, Innsbruck: Studienverlag, S. 63-79.
- Messick, S. (1995): *Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning*. American Psychologist 50 (9), S. 741-749.
- Nehring, A. (2014): *Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung*. Berlin: LOGOS (177).
- Padilla, M. J., McKenzie, D. L. & Shaw, E. L. (1986). *An examination of the line graphing ability of students in grades seven through twelve*. School Science and Mathematics 86, S. 20-26.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P., & Klopp, A. (2001). Naturwissenschaftliche Grundbildung. Testkonzeption und Ergebnisse. In Deutsches PISA-Konsortium (Hrsg.). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S.192-248). Opladen: Leske + Budrich.
- Scherer, R. (2014a): *Komplexes Problemlösen im Fach Chemie - Ein domänenspezifischer Zugang*. Zeitschrift für Pädagogische Psychologie 28 (4), S. 181–192.
- Scherer, R. (2014b): *Analyse der Struktur, Messvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. Eine Querschnittsstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*. Berlin: LOGOS (141).
- Schnotz, W. (1993): Wissenserwerb mit logischen Bildern. In B. Weidenmann (Ed.), *Wissens-erwerb mit Bildern: Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen*. Bern: Hans Huber, S. 95-147.
- Schnotz, W. (2001): *Wissenserwerb mit Multimedia*. Unterrichtswissenschaft, 29, S. 292-318.

- Schreiber, N., Theyßen H. & Schecker, H. (2009): *Experimentelle Kompetenz messen?! Physik und Didaktik in Schule und Hochschule*, PhyDid 3/8, S. 92 – 101.
- Skender, S. (2013): *Reliabilitäts- und Validitätsuntersuchungen zum Kognitiven Fähigkeitstest KFT 4-12+ R an einer Stichprobe von Fünft- und Siebtklässlern der Willy-Brandt-Gesamtschule München*. Dissertation, Philosophische Fakultät der Universität Rostock.
- Stern, E., Aprea, C. & Ebner, H. G. (2003). *Improving cross-content transfer in text processing by means of active graphical representation*. Learning and Instruction 13, S. 191-203.
- Süß, H.-M. (1996): *Intelligenz, Wissen und Problemlösen. Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen*. Göttingen: Hogrefe
- Vorholzer, A., Hägele, J. J. & von Aufschnaiter, C. (2017). *Entwicklung prozessbezogener Kompetenzen – eine videogestützte Analyse*. In: C. Maurer (Hg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis*. Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Zürich 2016, S. 288-291.
- Wavering, M. J. (1989). *Logical reasoning necessary to make line graphs*. Journal of Research in Science Teaching, 26, S. 373-379.
- Wenzel, D. (2016): *Problemlösestrategien und kognitive Fähigkeiten von Schülerinnen und Schülern im naturwissenschaftlichen Unterricht*. Masterarbeit, Institut für Fachdidaktik der Naturwissenschaften, Abt. Chemie und Chemiedidaktik, TU Braunschweig.
- Winn, W.D. (1987). *Charts, graphics and diagrams in educational materials*. In D. Willows and H. Houghton (Eds.), *The Psychology of Illustration*. Vol. 1. Basic Research. New York: Springer, S. 152-198.
- Wirth, J., Leutner, D. & Klieme, E. (2005). Problemlösekompetenz - Ökonomisch und zugleich differenziert erfassbar? In: E. Klieme, D. Leutner, J. Wirth (Hg.), *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 73-82.
- Zentrum für Internationale Bildungsvergleichsstudien (2013): *PISA 2012: Die Schülerinnen und Schüler in Deutschland überzeugen auch im Problemlöse-Test*. Pressemitteilung des Zentrums für Internationale Bildungsvergleichsstudien zur Veröffentlichung der Ergebnisse aus dem Kompetenzbereich Problemlösen in PISA 2012, Technische Universität München, [https://www.pisa.tum.de/fileadmin/w00bgi/www/Pressemitteilungen/PM\\_ZIB\\_PISA\\_2012\\_Problemloesen\\_final.pdf](https://www.pisa.tum.de/fileadmin/w00bgi/www/Pressemitteilungen/PM_ZIB_PISA_2012_Problemloesen_final.pdf), zuletzt aufgerufen am 05.11.2017.